JOURNAL OF COMPUTER AND SYSTEM SCIENCES 7, 506–521 (1973)

Distributions of Queue Lengths and Waiting Times in a Loop with Two-Way Traffic

ALAN G. KONHEIM*

IBM Research Center, Yorktown Heights, New York, U.S.A.

AND

BERND MEISTER

IBM Research Laboratory, Zurich, Switzerland

Received May 24, 1971

A loop transmission system consisting of a central station, N input terminals with infinite buffer, and a multiplexed channel is considered. Data flows from the terminals to the central station and from the central station to the terminals. The first type of data transfer is accorded priority over the second type. This necessitates intermediate buffering of data from the central station to the terminals. The distributions of the queue lengths at all terminals and of the virtual waiting times of data units are calculated. Two numerical examples are given.

1. INTRODUCTION

Loop or ring-switched transmission systems are of increasing interest in computer communications. A survey of recent results and an extensive bibliography can be found in [1]. Figure 1 shows a typical loop service system. N buffered input terminals $T^{(1)}, T^{(2)},..., T^{(N)}$ are linked by a common channel to a central station, e.g., a CPU. A loop service system with traffic only from the terminals to the CPU and with the "loop discipline" was analyzed in [2]. The loop discipline establishes priority service on the basis of position on the loop. Terminal $T^{(i)}$ is closer to the central station than terminal $T^{(j)}$ if $i < j$ and accordingly receives a higher grade of service. In [3] a loop system with two-way traffic was considered where the traffic from the CPU to the terminals was described in terms of the holding times of the channel by the CPU rather than explicitly.

* The research of this author was partially supported by the United States Air Force under Contract No. F 44620-70-C-0063.

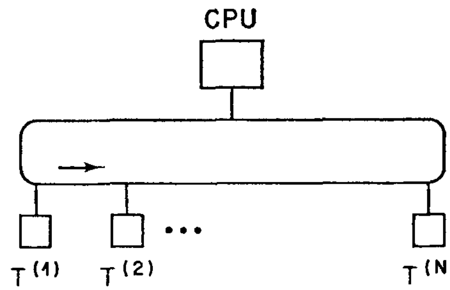

FIG. 1. Loop service system.

In this paper we are again concerned with the loop system of Fig. 1 and we now admit two service operations: the transmission of data from a terminal to the central station and from the central station to a terminal. The first type of service operation is accorded preemptive priority over the latter. Therefore, intermediate buffering of the traffic from the CPU has to be provided.

Performance will be measured in terms of (1) the buffering requirements as described by the stationary distributions of the queue lengths at the terminals, and (2) the response times, the stationary distributions of the waiting times, the times required to transmit a message from a terminal to the central station and from the central station to a terminal.

The common channel is time multiplexed. That is, the time axis $0 \leqslant t < \infty$ is divided into contiguous intervals or *slots* $s_j : (j-1)\Delta \leqslant t < j\Delta$ $(1 \leqslant j < \infty)$. Each slot is capable of transmitting a single *data unit* (with addressing information as required). What consitutes a data unit may vary in application. It may be a byte, a character or a fixed-size block. The channel plays the role of the server in our queueing model and the data units assume the role of customers.

We adhere to the following notational convention: subscripts will refer to evolution in time while superscripts will refer to position on the loop. We shall employ both the superscripts 0 and $N+1$ for the central station. The former when the central station is the source of a transmission and the latter when it is the destination of same.

In Sec. 2 we collect certain preliminary results which will be needed in the rest of the paper. Section 3 contains the description of the queueing model and the waiting line analysis. The waiting times are determined in Sec. 4 and several numerical examples are given in Sec. 5.

2. PRELIMINARIES

By *a standard process* $\mathscr{X}$ we shall mean a sequence $\mathscr{X} = \{X_j : 1 \leqslant j < \infty\}$ of independent, identically distributed, nonnegative, integer-valued random variables. The generating function of the process is the function $P(z) = E\{z^{X_1}\}$ defined for

$|z| \leqslant 1$ and analytic for $|z| < 1$. While we shall not assume a particular $P(z)$ in the analysis which follows, numerical results in Sec. 5 will use the compound Poisson process

$$P(z) = \exp \lambda \left(\frac{(1-q)z}{1-qz} - 1\right) \qquad (0 \leqslant q < 1; 0 \leqslant \lambda < \infty). \tag{2.1}$$

In our context, X_j is equal to the number of data units generated at a terminal or at the CPU during the jth slot. The particular choice of the law (2.1) is motivated by some recent measurements made of traffic in data channels [4].

Let $P(z)$ be a probability generating function and assume $\mu = P'(1) < 1$. According to Rouche's theorem [5] there exists a unique root in the unit disk, $\theta(P; w)$, of the equation

$$z - wP(z) = 0 \qquad (|w| < 1). \tag{2.2}$$

According to Lagrange's Theorem [6] $\theta(P; w)$ is analytic in $|w| < 1$. This root admits a probabilistic interpretation to which we now turn.

Let $\mathscr{X} = \{X_j : 1 \leqslant j < \infty\}$ be a standard process, $P(z) = E\{z^{X_1}\}$ and $\mu = E(X_1) < 1$. We consider the process

$$\mathscr{L} = \{L_j : 0 \leqslant j < \infty\}$$

defined by

$$L_j = (L_{j-1} - 1)^+ + X_j \qquad (1 \leqslant j < \infty) \tag{2.3}$$

with $a^+ = \max(a, 0)$. The *initial state* L_0 of $\mathscr{L}$ is a random variable (taking nonnegative integer values) with arbitrary law. The process (2.3) is a Markov chain with stationary transition law. Indeed if we set $H_j(z) = E\{z^{X_j}\}$ $(0 \leqslant j < \infty)$, then from (2.3)

$$H_j(z) = \frac{P(z)}{z} (H_{j-1}(z) + (z-1) H_{j-1}(0)) \qquad (1 \leqslant j < \infty). \tag{2.4}$$

If we write

$$H(z, w) = \sum_{j=0}^{\infty} H_j(z) w^j \qquad (|z|, |w| < 1),$$

then (2.4) yields

$$H(z, w) = \frac{H_0(z)z + w(z-1) P(z) H(0, w)}{z - wP(z)}. \tag{2.5}$$

The numerator of (2.5) contains the unknown boundary term $H(0, w)$. It is determined by an appeal to analyticity; since $H(z, w)$ is analytic in the polydisk $\{(z, w): |z| < 1, |w| < 1\}$ the vanishing of the denominator of (2.5) when $z = \theta(P; w)$ requires the same to occur in the numerator and thus

$$H(z, w) = \frac{H_0(z)z + w(z-1)\,P(z)\,\dfrac{H_0(\theta(P; w))}{1-\theta(P; w)}}{z - wP(z)}. \tag{2.6}$$

Some simplification is achieved when we restrict our attention to the stationary solution of (2.3). It is known [7] that $\{L_j\}$ converges in law, i.e., $H^*(z) = \text{limit } H_j(z)$ exists. The limit law may be obtained by first multplying (2.6) by $(1 - w)$ and then evaluating the limit as $w \to 1$, employing for this purpose a standard Tauberian theorem [8]. Carrying out the calculation we obtain:

LEMMA 2.1. *If* $L_j = (L_{j-1} - 1)^+ + X_j$ $(1 \leqslant j < \infty)$, $P(z) = E\{z^{X_j}\}$ *and* $\mu = E\{X_j\} < 1$, *then*

$$H^*(z) = \textit{limit } E\{z^{L_j}\} = (1-\mu)\frac{(z-1)\,P(z)}{z - P(z)}. \tag{2.7}$$

Remark. Equation (2.3) provided the starting point in [2] for the study of the loop system with traffic only from the terminals to the central station. The state variables $\{L_j\}$ count the total number of data units buffered at all terminals (at the start of the jth slot) while $\{X_j\}$ counts the total number of data units arriving in the system.

If $\mu = E\{X_1\}$, $\sigma^2 = \operatorname{Var}\{X_1\}$, and $\mu_3 = E\{(X_1 - \mu)^3\}$, then

$$E\{L^*\} = \frac{1}{2}\frac{\sigma^2}{1-\mu} + \frac{1}{2}\mu, \tag{2.8}$$

$$\operatorname{Var}\{L^*\} = \frac{\mu_3}{3(1-\mu)} + \left\{\frac{1}{2}\frac{\sigma^2}{1-\mu} + \frac{1}{2}(1-\mu)^2\right\}^2 - \frac{1}{12}(2\mu-1)(2\mu-3) \tag{2.9}$$

with $L^* = \text{limit } L_j$ (in law).

The process $\mathscr{L}$ has certain natural renewal points, the times $\tau_1 < \tau_2 < \cdots < \tau_n < \cdots$ at which $L_j = 0$. At these points the system renews itself in the sense that its behavior is independent of the past. It is clear that the random variables $\{\tau_j - \tau_{j-1} : 1 < j < \infty\}$ are independent and identically distributed. From (2.6) we obtain:

LEMMA 2.2.

$$E\{w^{(\tau_j - \tau_{j-1})}\} = \begin{cases} H_0(\theta(P; w)), & \textit{if } j = 1 \\ \theta(P; w), & \textit{if } 1 < j < \infty. \end{cases} \tag{2.10}$$

Next, we generalize the system of (2.3); suppose $\mathscr{X}^{(i)} = \{X_j^{(i)}: 1 \leqslant j < \infty\}$ are two independent standard processes, $P^{(i)}(z) = E\{z^{X_j^{(i)}}\}$ and $\mu^{(i)} = E\{X_j^{(i)}\}$ with $\mu^{(1)} + \mu^{(2)} < 1$. We consider the vector process

$$L_j^{(1)} = (L_{j-1}^{(1)} - 1)^+ + X_j^{(1)},$$
$$L_j^{(2)} = (L_{j-1}^{(2)} - (1 - L_{j-1}^{(1)})^+)^+ + X_j^{(2)} \qquad (1 \leqslant j < \infty). \tag{2.11}$$

The initial state $\mathbf{L}_0 = (L_0^{(1)}, L_0^{(2)})$ is a random variable with arbitrary distribution and we set $H_j^{(i)}(z) = E\{z^{L_j^{(i)}}\}$. According to Lemma 2.1 the first process $\{L_j^{(1)}\}$ converges in law. To study the second process we use Lemma 2.2 and introduce the notion of *a service epoch*. If $\tau_1 < \tau_2 < \cdots < \tau_n < \cdots$ are the times at which $L_j^{(1)} = 0$ then (2.11) implies

$$L_j^{(2)} = \begin{cases} L_{j-1}^{(2)} + X_j^{(2)} & \text{if } j - 1 \notin \{\tau_i\} \\ (L_{j-1}^{(2)} - 1)^+ + X_j^{(2)} & \text{if } j - 1 \in \{\tau_i\}. \end{cases} \tag{2.12}$$

In the context of a loop system with N terminals, $L_j^{(1)}$ will denote the total number of of data units buffered at terminals $T^{(1)},\ldots, T^{(k)}$, and $L_j^{(2)}$ the total number of data units buffered at terminals $T^{(k+1)},\ldots, T^{(N)}$. The jth slot is a service epoch for this second collection of terminals if and only if $L_{j-1}^{(1)} = 0$, i.e., if and only if $j - 1 \in \{\tau_i\}$. When we write (2.12) in the form

$$L_{\tau_j}^{(2)} = (L_{\tau_{j-1}}^{(2)} - 1)^+ + \sum_{k=1+\tau_{j-1}}^{\tau_j} X_k^{(2)} \tag{2.13}$$

we recognize its similarity to (2.3). The independence and identical distribution of $\{\tau_j - \tau_{j-1} : 1 < j < \infty\}$ implies that the random variables

$$\sum_{k=1+\tau_{j-1}}^{\tau_j} X_k^{(2)}$$

consitute a standard process and hence (2.13) is a special case of (2.3) with $P(z)$ replaced by $\theta(P^{(1)}(z); P^{(2)}(z))$; the latter is the generating function of the effective input process in (2.13). From (2.2) we easily see that the expected number of arrivals (over one renewal cycle) is just $\mu^{(2)}/1 - \mu^{(1)}$ which is less than one if $\mu^{(1)} + \mu^{(2)} < 1$. We thus have:

LEMMA 2.3. *If $\mu^{(1)} + \mu^{(2)} < 1$, then limit*

$$E\{z^{L_j^{(2)}}\} = \frac{1 - \mu^{(1)} - \mu^{(2)}}{1 - \mu^{(1)}} \frac{(z - 1)\,\theta(P^{(1)}(z); P^{(2)}(z))}{z - \theta(P^{(1)}(z); P^{(2)}(z))}. \tag{2.14}$$

Remark. The system (2.11) is equivalent to the system

$$
\begin{aligned}
L_j^{(1)} &= (L_{j-1}^{(1)} - 1)^+ + X_j^{(1)}, \\
L_j^{(1)} + L_j^{(2)} &= (L_{j-1}^{(1)} + L_{j-1}^{(2)} - 1)^+ + X_j^{(1)} + X_j^{(2)}.
\end{aligned} \tag{2.15}
$$

3. The Queueing Model and the Stationary Queue Length

We shall model the traffic to and from the central station by standard processes. We shall use both the superscripts 0 and $N + 1$ in referring to the central station; the former when the central station is the source of the data and the latter when it is the destination. Thus

$$\mathscr{X}^{(i;N+1)} = \{X_j^{(i;N+1)}: 1 \leqslant j < \infty\} \qquad (1 \leqslant i \leqslant N)$$

defines the process by which data enters at $T^{(i)}$ for transmission to the central station, while

$$\mathscr{X}^{(0;i)} = \{X_j^{(0;i)}: 1 \leqslant j < \infty\} \qquad (1 \leqslant i \leqslant N)$$

defines the process by which data is generated at the central station for transmission to $T^{(i)}$. We assume that the $2N$ processes are independent and we set $P^{(i;N+1)}(z) = E\{z^{X_j^{(i;N+1)}}\}$ and $P^{(0;i)}(z) = E\{z^{X_j^{(0;i)}}\}$. It will be convenient to introduce the following convention with superscripts; if $\alpha \subseteq \{0, 1, 2,..., N\}$ and $\beta \subseteq \{1, 2,..., N+1\}$ then $X_j^{(\alpha;\beta)}$ is equal to the total number of data units arriving in the system at one of the stations $T^{(k)}$ with $k \in \alpha$ for transmission to one of the stations $T^{(k')}$ with $k' \in \beta$, where a station is either terminal or the central station.

Definition 3.1. If $\alpha \subseteq \{0, 1, 2,..., N\}$ and $\beta \subseteq \{1, 2,..., N + 1\}$, the $L_j^{(\alpha;\beta)}$ is the total number of data units buffered at the stations $T^{(k)}$ $(k \in \alpha)$ for transmission to one of the stations $T^{(k')}$ $(k' \in \beta)$ just before the start of the $(j + 1)^{\text{st}}$ slot, i.e., at time $j\Delta - 0$.

Remark. $L_j^{(\alpha;\beta)}$ refers to the state of the system at different times. The jth slot "starts" at the central station and "moves" along the loop passing each of the terminals $T^{(1)}, T^{(2)},..., T^{(N)}$ in turn. The state of the buffer at $T^{(k)}$ at the start of the jth slot will refer to the state at the arrival of this slot. This presents no difficulties.

It remains to describe the queue discipline—how the channel is allocated among the $N + 1$ users. Our queue discipline will accord highest priority to traffic *from a terminal to the central station.* When the terminals $T^{(i_1)}, T^{(i_2)},..., T^{(i_k)}$ $(1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant N)$ jointly request a slot, the terminal "closest" to the central station $T^{(i_1)}$ will claim the slot. This is the loop discipline of [2]. It may be helpful to imagine a server (the slot) who makes repeated tours of the terminals on the loop visiting them in the order $T^{(1)}, T^{(2)}, \ldots, T^{(N)}$. On each tour the server may bring at most one data unit (customer) from one of these terminals back to the central station. This data unit is taken from the first terminal (in the order of his visit) seeking such a transfer. Continuing with this analogy, the server may on occasion depart from the central station—the origin of his tour—with a data unit destined for one of the N terminals. This traffic *from the central station to a terminal* is of lower priority. The transmission of such a data unit d whose destination is terminal $T^{(k)}$ will be interrupted if upon arrival at terminal $T^{(i)}$ $i < k$, the buffer at this terminal contains either:

(1) a data unit for transmission to the central station, or

(2) a data unit for transmission to $T^{(k')}$ with $k' \geqslant k$.

In either case the data unit d has its "journey" interrupted and is temporarily buffered at $T^{(i)}$ where it will compete with other data units for the use of subsequent slots. It is convenient to imagine that at $T^{(i)}$ there develop $N - i + 1$ queues $Q^{(i;k)}$ $(i < k \leqslant N + 1)$ for the data units buffered at $T^{(i)}$ with destination $T^{(k)}$ $(i < k \leqslant N + 1)$. The queue discipline is specified in Table I.

Remarks. (i) Table I will also be used to prescribe how the jth slot is assigned by the central station to transmit data to $\{T^{(i)}: 1 \leqslant i \leqslant N\}$ [Cases 1(a) and 1(b)].

(ii) Case 3 implies that data arriving at a particular terminal is processed in the order of arrival.

The main lemma is:

LEMMA 3.2. *If* $i \leqslant k \leqslant N$,

$$L_j^{(0,1,\ldots,i;k+1,\ldots,N+1)} = (L_{j-1}^{(0,1,\ldots,i;k+1,\ldots,N+1)} - 1)^+ + \sum_{s=k+1}^{N} X_j^{(0;s)} + \sum_{s=1}^{i} X_j^{(s;N+1)}. \tag{3.1}$$

Proof. The proof is by induction; hence we assume

$$L_j^{(0,1,\ldots,i-1;k+1,\ldots,N+1)} = (L_{j-1}^{(0,1,\ldots,i-1;k+1,\ldots,N+1)} - 1)^+ + \sum_{s=k+1}^{N} X_j^{(0;s)} + \sum_{s=1}^{i-1} X_j^{(s;N+1)} \tag{3.2}$$

and we shall prove

$$L_j^{(i;k+1,\ldots,N+1)} = (L_{j-1}^{(i;k+1,\ldots,N+1)} - (1 - L_{j-1}^{(0,1,\ldots,i-1;k+1,\ldots,N+1)})^+)^+ + X_j^{(i;N+1)}. \tag{3.3}$$

Equations (3.2) und (3.3) together imply (3.1) (see the remark following Lemma 2.3).

TABLE I

	State of the jth slot	State of the buffer at $T^{(i)}$	Action taken
1(a)	Free	$Q^{(i;k)} = \varnothing \quad (n < k \leqslant N+1)$ $Q^{(i;n)} \neq \varnothing$	jth slot is acquired by the data unit at the head of the queue $Q^{(i;n)}$
1(b)		$Q^{(i;k)} \neq \varnothing \quad (i < k \leqslant N+1)$	jth slot remains free
2(a)	Contains a data unit d for $T^{(i)}$	$Q^{(i;k)} = \varnothing \quad (n < k \leqslant N+1)$ $Q^{(i;n)} \neq \varnothing$	d arrives at its destination and the slot is acquired by the data unit at the head of the queue $Q^{(i;n)}$
2(b)		$Q^{(i;k)} = \varnothing \quad (i < k \leqslant N+1)$	d arrives at its destination and the jth slot is free
3	Contains a data unit d for $T^{(r)}$ $(i < r \leqslant N)$		d joints the end of the queue $Q^{(i;r)}$, the slot is made free and assigned as in 1(a)
4	Contains a data unit for the central station		The data unit retains the slot

There are two cases to consider:

Case (*a*). $L_{j-1}^{(0,1,\ldots,i-1,;k+1,\ldots,N+1)} = 0$.

In this case the jth slot is either free or contains a data unit for one of the terminals $T^{(m)}$ with $i \leqslant m \leqslant k$. In any event the slot is made available to effect a transfer of a data unit from $T^{(i)}$ to $T^{(m)}$ with $k+1 \leqslant m \leqslant N+1$. Thus

$$L_j^{(i;k+1,\ldots,N+1)} = (L_{j-1}^{(i;k+1,\ldots,N+1)} - 1)^+ + X_j^{(i;N+1)}$$

which is in agreement with (3.3) in this case.

Case (*b*). $L_{j-1}^{(0,1,\ldots,i-1;k+1,\ldots,N+1)} > 0$.

The jth slot contains a data unit for one of the terminals $T^{(k+1)},\ldots, T^{(N+1)}$ and is assigned at $T^{(i)}$ according to the rules of Cases 3 and 4. Either the slot is retained by the data unit presently holding it (as in Case (4) or there is an exchange made. In both cases there is no net change in the number of data units buffered at $T^{(i)}$ and hence

$$L_j^{(i;k+1,\ldots,N+1)} = L_{j-1}^{(i;k+1,\ldots,N+1)} + X_j^{(i;N+1)}$$

which is in agreement with (3.3).

DEFINITION 3.3. *The* jth *slot is a service epoch for* $T^{(i)}$ if this slot upon its arrival at $T^{(i)}$ is either free or contains a data unit for $T^{(i)}$. Let $\{\tau_j^{(i)}: 1 \leqslant j < \infty\}$ denote the service epochs for $T^{(i)}$.

Since the service epochs for $T^{(i)}$ are the times at which $L_{j-1}^{(0,1,\ldots,i-1;i+1,\ldots,N+1)} = 0$ we have from Lemma 2.2:

LEMMA 3.4.

(1) *The random variables* $\{\tau_j^{(i)} - \tau_{j-1}^{(i)} : 1 < j < \infty\}$ *are independent and identically distributed.*

(2) $E\{w^{(\tau_j^{(i)}-\tau_{j-1}^{(i)})}\} = \theta(P; w)$ $(1 < j\ \infty)$ *where*

$$P(z) = \prod_{s=i+1}^{N} P^{(0;s)}(z) \prod_{s=1}^{i-1} P^{(s;N+1)}(z). \tag{3.4}$$

Lemmata 3.2, 3.4, and 2.3 then imply:

THEOREM 3.5. *If* $\sum_{s=i+1}^{N} \mu^{(0;s)} + \sum_{s=1}^{i} \mu^{(s;N+1)} < 1$, *then*

$$E\{z^{L^{(i;i+1,\ldots,N+1)}_{\tau_j^{(i)}}}\} \to E\{z^{L^{(i;i+1,\ldots,N+1)^*}}\}$$
$$= \frac{1 - \sum_{s=i+1}^{N} \mu^{(0,s)} - \sum_{s=1}^{i} \mu^{(s,N+1)}}{1 - \sum_{s=i+1}^{N} \mu^{(0,s)} - \sum_{s=1}^{i-1} \mu^{(s,N+1)}} \; \frac{(z-1)\, \theta(P; P^{(i;N+1)})}{z - \theta(P; P^{(i;N+1)})} \tag{3.5}$$

with P *given by* (3.4).

4. THE STATIONARY DELAY

There are two response times of interest in this loop system; the delay in transmitting a message from the central station to terminal $T^{(i)}$ and the delay in transmitting a message from terminal $T^{(i)}$ to the central station. The latter delay has already been calculated ([2], [9]). Indeed the data flow from terminals to the central station is not impeded by the data flow from the central station to the terminals and hence as far as the former flow is concerned, this loop system is equivalent to one in which there is only flow from the terminals. To calculate the delay in transmitting a message of M data units from the central station to terminal

$T^{(i)}$ we assume that this message is available at time $j\Delta + 0$. At this time there are

$$(L_j^{(0,1,\ldots,i-1;i,\ldots,N+1)} - 1)^+$$

data units in the system with higher priority of service. During the interval $j\Delta < t \leqslant (j+1)\Delta$ an additional

$$\sum_{s=i+1}^{N} X_{j+1}^{(0;s)} + \sum_{s=1}^{i-1} X_{j+1}^{(s;N+1)}$$

data units will arrive which have higher priority of service. Now consider the system (2.3),

$$L_k = (L_{k-1} - 1)^+ + X_k (1 \leqslant k < \infty) \tag{4.1}$$

with

$$L_0 = M + (L_j^{(0,1,\ldots,i-1;i,\ldots,N+1)} - 1)^+ + \sum_{s=i+1}^{N} X_{j+1}^{(0;s)} + \sum_{s=1}^{i-1} X_{j+1}^{(s;N+1)}, \tag{4.2}$$

$$X_k = \sum_{s=i+1}^{N} X_{j+k}^{(0;s)} + \sum_{s=1}^{i-1} X_{j+k}^{(s;N+1)} \qquad (1 \leqslant k < \infty). \tag{4.3}$$

The first time τ at which $L_\tau = 0$ is the time at which the service operation—transmission of these M data units to $T^{(i)}$—will be completed. These M data units will have then spent $\tau + 1$ units of time within the system and thus experience a delay of $\tau + 1 - M$ time units. The delay $D_j^{(i)}(M)$ is thus given by Lemma 2.2. If we set $D^{(i)}(M)^* = \text{limit } D_j^{(i)}(M)$ (in law) we have:

THEOREM 4.1. *If* $\sum_{s=i}^{N} \mu^{(0;s)} + \sum_{s=1}^{i-1} \mu^{(s;N+1)} < 1$

$$E\{w^{D^{(i)}(M)^*}\} = w \left\{\frac{\theta(P; w)}{w}\right\}^M (\theta(P; w) - 1) \left(1 - \sum_{s=i}^{N} \mu^{(0;s)} - \sum_{s=1}^{i-1} \mu^{(s;N+1)}\right) \frac{1}{w - P^{(0;i)}(\theta(P; w))} \tag{4.4}$$

with

$$P(z) = \prod_{s=i+1}^{N} P^{(0;s)}(z) \prod_{s=1}^{i-1} P^{(s;N+1)}(z). \tag{4.5}$$

5. NUMERICAL EXAMPLES

In this section we shall indicate how the development of Sections 2–4 may be used to obtain quantitative results for system performance.

The utilization (of the channel) at terminal $T^{(i)}$, denoted by $\rho^{(i)}$, is defined as the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_{\{L_j^{(0,1,\ldots,i;i+1,\ldots,N+1)}>0\}} . \tag{5.1}$$

Note that the ratio

$$\frac{1}{n} \sum_{j=0}^{n-1} \chi_{(L_j^{(0,1,\ldots,i;i+1,\ldots,N+1)}>0)} \tag{5.2}$$

is the fraction of the first n slots which contain a data unit upon arrival at the $(i+1)$st terminal. The utilization $\rho^{(i)}$ is the limiting value of the ratio (5.2) and is readily evaluated as

$$\rho^{(i)} = \sum_{k=i+1}^{N} \mu^{(0;k)} + \sum_{k=1}^{i} \mu^{(k;N+1)}, \tag{5.3}$$

with $\mu^{(0;k)} = E\{X_j^{(0;k)}\}$ and $\mu^{(k;N+1)} = E\{X_j^{(k;N+1)}\}$. Note that (5.1)–(5.3) have meaning when $i = 0$. The *channel utilization* is the maximum of these utilizations,

$$\rho = \max_{0\leqslant i\leqslant N} \rho^{(i)}.$$

Note that the sequence $\{\rho^{(i)}: 0 \leqslant i \leqslant N\}$ need not be monotonic. A simplification occurs when we assume that the processes $\{\mathscr{X}^{(i,N+1)}, 1 \leqslant i \leqslant N\}$ have all the same distribution and that also the processes $\{\mathscr{X}^{(0,i)}, 1 \leqslant i \leqslant N\}$ are identically distributed but with a different distribution. We set

$$\begin{aligned} P_1(z) &= E\{z^{X_j^{(i;N+1)}}\}, \qquad & \mu_1 &= E\{X_j^{(i;N+1)}\}, \\ P_2(z) &= E\{z^{X_j^{(0;i)}}\}, \qquad & \mu_2 &= E\{X_j^{(0;i)}\}. \end{aligned} \tag{5.4}$$

In this case $\rho^{(i)} = (N - i)\,\mu_2 + i\mu_1$ and

$$\rho = N \max(\mu_1 , \mu_2).$$

We may identify

$$\rho_{\text{CPU}} = N\mu_2 \qquad \rho_T = N\mu_1$$

as the utilization of the channel by the CPU and terminals, respectively. For illustrative purposes we shall make this assumption concerning the input processes and take

$$P_i(z) = \exp \gamma_i \left(\frac{(1-q_i)z}{1-q_iz} - 1\right) \qquad (i = 1, 2).$$

Our two examples will be

Example	γ_1	q_1	γ_2	q_2	N	ρ_{CPU}	ρ_T	ρ
1	0.02500	0.50	0.00625	0.75	10	0.50	0.25	0.50
2	0.00625	0.75	0.02500	0.50	10	0.25	0.50	0.50

In the first example more data flows from the terminals to the CPU than from the CPU to the terminals; in the second example the date flow is reversed.

The calculation of the stationary state probabilities

$$\Pr\{L^{(i;i+1,\ldots,N+1)} = k\} \qquad (0 \leqslant k < \infty) \tag{5.5}$$

requires the determination of the coefficients of the generating function

$$\mathscr{H}^{(i;i+1,\ldots,N+1)*}(z) = \frac{1-(N-i)\,\mu_2 - i\mu_1}{1-(N-i)\,\mu_2-(i-1)\,\mu_1}\,\frac{(z-1)\,\theta(P;P_1(z))}{z-\theta(P;P_1(z))} \tag{5.6}$$

with $P(z) = P_1^{i-1}(z)\,P_2^{N-i}(z)$. For the processes of (5.4) this inversion in closed form is possible although the resulting formulae are not amenable to direct inspection and interpretation. To obtain numerical results we shall employ the following device; suppose X and Y are independent random variables, $Q(z) = E\{z^X\}$, $H(z) = E\{z^Y\}$ and

$$Y \sim (Y-1)^+ + X \tag{5.7}$$

where $\sim$ indicates that the two sides agree in distribution. In this case

$$H(z) = (1-\mu)\,\frac{(z-1)\,Q(z)}{z-Q(z)} \qquad (\mu = E\{X\})$$

which has the form of (5.6). From (5.7) we obtain the recurrence

$$\Pr\{Y=n\} = \frac{\Pr\{Y=n-1\} - \sum_{j=1}^{n-1}\Pr\{Y=j\}\Pr\{X=n-j\} - \Pr\{Y=0\}\Pr\{X=n-1\}}{\Pr\{X=0\}} \tag{5.8}$$

which may then be employed to recursively calculate the state probabilities $\{\Pr\{Y=n\}: 0 \leqslant n < \infty\}$. The state probabilities (5.5) have been calculated by first determining the coefficients of the generating function $\theta(P;\, P_1(z))$ and then applying the recurrence (5.8). The results for examples 1 and 2 are given graphically in Figs. 2 and 3, respectively.

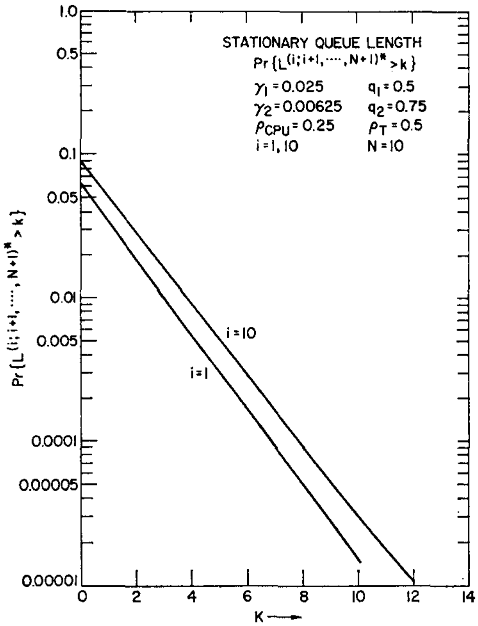

FIGURE 2

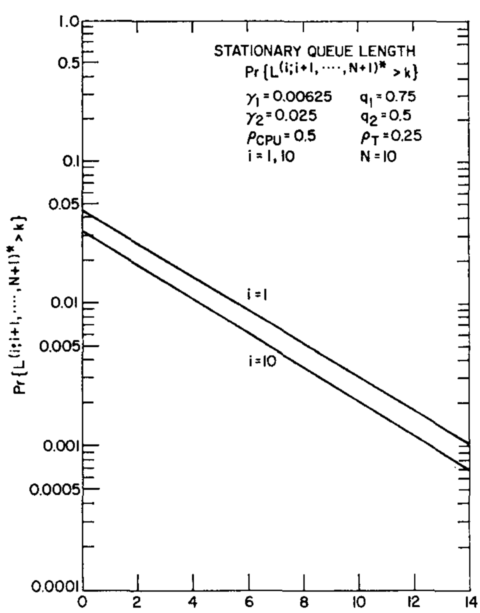

FIGURE 3

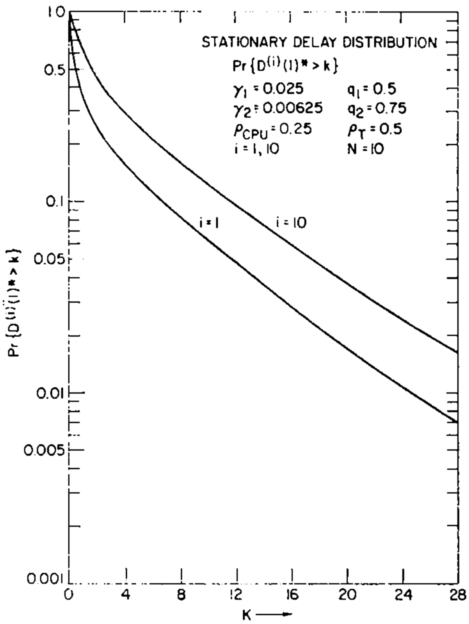

FIGURE 4

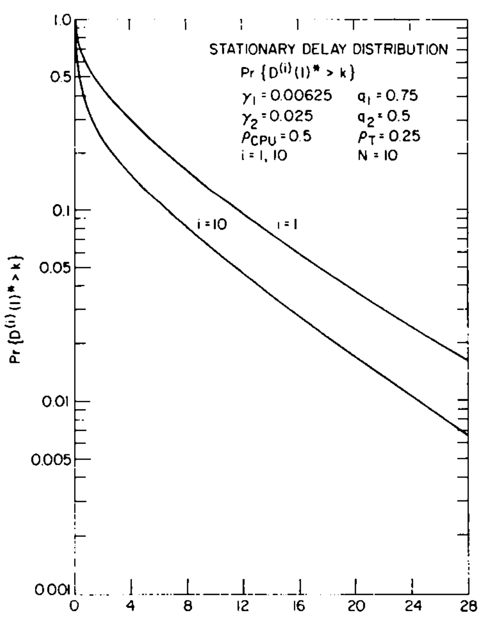

FIGURE 5

The stationary delay probabilities

$$\Pr\{D^{(i)}(M)^* = k\} \qquad (0 \leqslant k < \infty) \tag{5.9}$$

are the coefficients of the generating function

$$D_M^{(i)}(w)^* = (1 - \rho^{(i-1)})\, w \left(\frac{\theta(P; w)}{w}\right)^M \frac{\theta(P; w) - 1}{w - P_2(\theta(P; w))} \tag{5.10}$$

with P given as before. To obtain the coefficients of (5.10) we first write $D_M^{(i)}(w)^*$ as the product

$$D_M^{(i)}(w)^* = (1 - \rho^{(i-1)} + \mu_2)\, w \left(\frac{\theta(P; w)}{w}\right)^M (1 - \theta(P; w)) \cdot A(w)$$

with

$$A(w) = \frac{1 - \rho^{(i-1)}}{1 - \rho^{(i-1)} + \mu_2} \frac{1}{P_2(\theta(P; w)) - w}\,.$$

Next, we observe that with Y given by (5.7) we have

$$\sum_{k=0}^{\infty} \Pr\{Y \leqslant n\}\, z^n = (1 - \mu) \frac{1}{Q(z) - z}$$

so that the coefficients of $A(w)$ may be obtained from the recursion (5.8) provided we set $Q = P_2(\theta(P; z))$. The resulting sequence is then convolved with the sequence of coefficients of

$$(1 - \rho^{(i-1)} + \mu_2)\, w \left(\frac{\theta(P; w)}{w}\right)^M (1 - \theta(P; w))$$

to obtain the stationary delay probabilities. The results are portrayed for examples 1 and 2 in Figs. 4 and 5, respectively.

All the numerical results show the following: If more traffic flows from the terminals to the central station than into the opposite direction the terminals at the "beginning of the loop" experience a better grade of service than the terminals at the "end of the loop." If the flow pattern is reversed and more traffic is directed from the central station to the terminals many data units will be preempted and intermediately buffered at terminals which are close to the central station. Therefore, the terminals at the "end of the loop" now receive a higher grade of service.

Acknowledgment

The authors are indebted to Dr. John Cocke, IBM Fellow, for suggesting this problem.

REFERENCES

1. W. W. CHU AND A. G. KONHEIM, On the analysis and modeling of a class of computer communication systems, *IEEE Trans. Comm.* **COM-20** (1972), 645–660.
2. A. G. KONHEIM AND B. MEISTER, Service in a loop system, *J. Assoc. Comput. Mach.* **19** (1972), 92–108.
3. A. G. KONHEIM AND B. MEISTER, Two-way traffic in loop service systems, *Networks* **1** (1972), 291–301.
4. E. FUCHS AND P. E. JACKSON, Estimates of distributions of random variables for certain computer communications traffic models, *Commun. ACM* **13** (1970), 752–757.
5. L. V. ALHFORS, "Complex Analysis," p. 124, McGraw-Hill Book Co., New York, 1954.
6. E. T. WHITTAKER AND G. N. WATSON, "A Course of Modern Analysis," pp. 132–133, Cambridge University Press, Cambridge, 1952.
7. D. V. LINDLEY, The theory of queues with a single server, *Proc. Cambridge Phil. Soc.* **48** (1952), 277–289.
8. E. W. HOBSON, "Theory of Functions of a Real Variable," Cambridge University Press, Cambridge, 1927.
9. A. G. KONHEIM, "Service epochs in a loop system," Twenty-Second International Symposium on Computer Communication, Networks and Teletraffic, Polytech. Inst. Brooklyn, Brooklyn, New York, April 1972.